A STUDY OF CERTAIN GENERAL FACTORS PREDICTING SUCCESS AT PRODUCTION WORK

by

BENTLEY AUGUSTINE BARNABAS
A. B., University of Wichita, 1929

A THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Psychology

KANSAS STATE COLLEGE
OF AGRICULTURE AND APPLIED SCIENCE

1953

TABLE OF CONTENTS

INTRODUCTION

Purpose of the Study

Following the establishing of plant-wide standards for aptitude tests to be used in selecting workmen in the production plant of The Coleman Company, Inc., Wichita, Kansas, in May of 1944, the author saw the need for further follow-up work to determine what factors, if any, common to all good workmen, would distinguish them from those persons whose measured traits did not suit them to production work. The May, 1944 study revealed the fact that there were significant differences in the aptitudes needed in the different departments and that preliminary departmental standards could be set up to aid in placement.

Research into the general factors was needed to further refine these studies in the interest of finding out, (1) What critical scores might exist which would show up clearly those who could not succeed at production work, (2) What measurable traits the good workmen had in common as differentiated from measurable traits held in common by the poor workmen, and (3) If such a study might contribute to general knowledge of the field so that guidance workers and personnel placement workers could make more effective decisions in advising those who are seeking to enter production employment.

Description of the Coleman Company, Incorporated

This company provided an ideal locus for a study of this type since their manufacturing includes the fabrication of both large and small products ranging from furnaces and space heaters to small lanterns and camp stoves. At the time of the studies they employed approximately 1400 persons of whom approximately 75 per cent were in production work and production administration. The author was serving as a technical personnel consultant to this company throughout these studies in his capacity as the executive head of Associated Personnel Technicians. Both of these firms are located in Wichita, Kansas.

MATERIALS AND METHODS

Characteristics of the Sample

Selection of the Sample. Beginning in the late summer of 1944, foremen in the Coleman plant were asked to rate their workers on a five-point rating scale in which a rating of "1" was "best" and a rating of "5" was "worst". While worker production was known in some of the departments where pay was on an incentive basis, the ratings in the remaining departments were estimates based on the observation of workers by the foremen and supervisors. Three such ratings were obtained at three-month intervals.

From the ratings the author placed in a separate file all workers who earned a rating of "2" or "1" after training and referred to them as "good workmen". In another file the author placed all workers rated "4" or "5" and referred to them as "poor workmen".

All workers placed in the "good workmen" file were able to maintain their ratings for 5 months or more (at least two ratings). All workmen placed in the "poor workmen" file either earned ratings of "4" or "5" for the full period, were terminated at a time when they had such a rating or lapsed to a rating of "4" or "5" after an initial rating of "3".

The sample was reduced in number by the removal of those about whom there was doubt as to the validity of the rating. Those rated "3", "2", or "1" who terminated before a second rating could be obtained, those who improved to a rating of "3", but who terminated before it could be determined whether or not they might have further improved to a rating of "2", as well as workers who were promoted to a supervisory capacity before a second rating could be obtained, were all eliminated from this contrasting sample.

Age, Education, and Sex of the Workers in the Sample. The 103 "good workmen" had been with The Coleman Company, Inc., an average of 33.2 months, averaged 34.5 years in age, and had an average educational level equivalent to 10 school grades (average 10.4 school grades). In the group were 30 women production workers and 73 male production workers.

The "poor workmen" had been with the company an average of 24.8 months, averaged 34.9 years in age, and had an average educational level equivalent to 9 school grades (average of 9.1). This sample contained 25 women production workers and 79 male production workers, a total of 104.

Representativeness of the Sample. Both groups were representative of all major production departments and the numbers in both samples are quite proportionate with respect to the type of work done.

Types of work included: machine operation, mantle production, buffing and plating, tool and die making, stamp and punch press operation, small assembly, final assembly, tank assembly, re-work, inspection, maintenance, and decalcomania production. Proportionate representation in the sample was within 1 per cent of being exactly the same in all departments except 2. The decalcomania department had 3 more workers in the "poor" group than in the "good" group, while there were 4 more inspectors in the "good" group than in the "poor" group. Eight of the departments had identical representation. Actually the most significant differences in the representativeness of the sample were the excess of 5 more women workers in the "good" group than in the "poor" group and the additional 1.3 years of mean education shown by the "good" group.

The Test Battery

All of the workmen took the same tests under the same circumstances. All were administered by psychometrists of the staff of Associated Personnel Technicians and all took the tests after employment.

The battery included The Personnel Test, a 12-minute spiral omnibus test of mental efficiency consisting of 50 problems, The Mechanical Comprehension Test, The Personality Inventory, and The MacQuarrie Test for Mechanical Ability, a test comprising a series of seven sub-tests (more fully described below).

This study includes data derived only from The MacQuarrie series and The Personnel Test for the reason that The Personality Inventory and The Mechanical Comprehension Test scores are considerably different for women than for men. The author wished to include women production workers in this study and to investigate mental and psycho-motor functions rather than differences in personality or achievement.

The MacQuarrie Test. This series of tests is one of the oldest of the paper-and-pencil tests of psycho-motor coordination and visualization, containing seven subtests for each of which there is a practice test to be given in advance of each recorded test. It is a timed test emphasizing speed and accuracy. The method of administration by Associated Personnel Technicians assured that each subject understood the instructions, since the practice test was criticized by a monitor who inspected the work

of the subjects after the practice test was completed. Trained in this work, the monitors checked to see that the subject understood what was wanted and, if in doubt, waited while the subject attempted another item or two of the practice test under the watchful eye and supervision of the monitor. There was at least one monitor for every ten subjects taking the test in group situations. The seven MacQuarrie subtests sample the subject's performance in the following manner and order:

Test 1 requires the subject to make a continuous curving line from left to right and then from right to left on successive rows, the sharp point of the #2½ pencil passing through fine openings in vertical lines blocking the path of the pencil and requiring curving motion. Touching the vertical lines with the pencil is an error and the scorers must be able to see white space between the pencil mark and the sides of the opening for the passage to count. Openings are 3/64 of an inch wide.

Test 2 is a simple tapping test of broad tolerance requiring high-speed rhythmic action of the forearm. Little accuracy is required, the subject receiving a count for every dot he makes in each of a series of circles up to three dots in each circle. So long as the dots are within the circle (so as to be seen therein by the scorer) the dot counts. Circles are 11/32 of an inch in diameter.

Test 3 has smaller circles along a continuous line, the line between the dots arranged so that the subject follows it along rows first from left to right and then right to left. The sub-

ject is asked to make one dot in the small circle and the dot may not touch the circumference of the circle at any point if it is to count. The circles are spaced at different distances throughout the test so that the subject must make accurate aim with his whole arm as he progresses to a new circle and stops to make the dot. Circles are 11/64 of an inch in diameter and range from a minimum separating distance of 1/32 of an inch to a maximum of 1 3/4 inches apart.

Test 4 begins the subtests that stress visualization. This test asks the subject to copy an angular diagram in a square to the side of the diagram shown, using dots formed in a grid as a reference point for the copying.

Test 5 introduces rather ingeniously a slight element of the abstract to the operation described above in Test 4. Here the subject is expected to locate in a series of small squares, certain dots which stand in the position of letters of the alphabet which are in a large square on the center of the page. The change in size is the principal problem for the subject in this test, since the large square containing the letters is four times the size of the smaller squares with the dots.

Test 6 is a blocks test in which the subject is expected to "see" how many blocks in a pile touch certain blocks with X's marked on them. Obviously it is a test of visualization of three-dimensional spatial relations since the blocks are simply perspective drawings with some of the oblong blocks standing on end, others lying flat in perspective.

Test 7 is a maze of curving lines. Each line starts from one of ten numbered squares at the left and ends in a square at the right. The subject places in the boxes at the right side of the maze (where the lines end) the number of the square in which the line started at the left. The test places a premium on accurate vision and acuity.

The Personnel Test. The Personnel Test has already been described as a spiral omnibus problem-solving test of mental efficiency administered in 12 minutes. It was derived from the Otis type test by E. F. Wonderlic with certain improvements in order of difficulty and arrangement.

Statistical Treatment

The author calculated the means and standard deviations on all scales for both groups, calculated the significance of the differences in means, figured the bi-serial coefficients of correlation where desirable, constructed a correlation matrix for both groups, and completed a factor analysis using the Complete Centroid Method.[1]

---

[1] L. L. Thurstone, "Multiple Factor Analysis", p. 161 et sequ.

THE FINDINGS OF THE STUDY

Differences in Mean Scores

Significance of the Difference. The "good workmen" group had higher mean scores on each of the subtests in the battery. Significant differences between the two groups ranged from quotients of 2.81 to 4.54 when the actual differences in the mean scores were divided by the standard deviations of the differences between the means. Table 1 shows these quotients and the chances in 1000 that extended sampling would continue to produce a difference greater than zero.

Table 1. Significance of difference in mean score.

Test or subtest	Quotient of actual difference over standard error of difference	Chances in 1000 that the difference is greater than zero
MacQuarrie #1	2.81	997
MacQuarrie #2	3.54	999+
MacQuarrie #3	4.54	999+
MacQuarrie #4	3.79	999+
MacQuarrie #5	3.13	999+
MacQuarrie #6	3.88	999+
MacQuarrie #7	4.48	999+
Personnel Test	3.76	999+

From Table 1 it is evident that the differences in the two groups were most significant in Test 3 and Test 7 of the Mac-Quarrie and least significant in Test 1. In appraising the factor analysis to be described later, these significant differences should be kept in mind, especially the fact that 4 tests outranked the Personnel Test in significance of the differences.

These differences, in themselves, indicate that the series of tests provide a good tool for the selection of production workers.

Characteristics of the Distributions. Visual inspection revealed that the distributions were quite normal in the "good workmen" group. Two distributions in the "poor workmen" group deviated slightly from the classic normal distribution. The distribution on Test 4 in the poorer group was clearly bi-modal with the mean of the group occurring almost exactly equidistant between the two modes. It is to be noted that this did not seriously affect the linearity of the regression in the correlations of these scores with other tests, although it certainly served to increase slightly the coefficients of correlation between this test and others.

The other distribution which deviated from normal was that for Test 6 in the "poor workmen" group where there were only 1.45 standard deviations below the mean in the range of scores and 2.2 S.D.'s above. The reason for this definite skew lies in the fact that 9 per cent of the "poor workmen" earned the score of zero, failing to answer even one item correctly. Better

measurement, of course, would be obtained if the test were not quite so difficult and had a lower zero point, although in its present form it still discriminates clearly between those who can "see" this type of perspective and those who cannot. It is useful for selection in tasks where the test has been validated, but would be far less useful in guidance or other more critical work.

Critical Scores

"Breaking Points" in the Distributions. In addition to the differences in the means, critical scores were discovered in the study of these distributions. In each test there was a clear-cut breaking point between the "good workmen" and the "poor workmen". Table 2 shows how well these breaking points served to predict that any person below the point was predictably a poor risk at production work of any kind.

Table 2 points up the discriminative power of a test of this type when properly validated. The 115 scores below the "critical" score by "poor workmen" were made by 42 persons, while the 8 scores made below these critical scores by "good workmen" were made by 3 persons. Of the poor workmen, 14 went below the critical score once, 9 went below twice, 6 went below 3 times, 5 went below 4 times, 4 went below 5 times, and 4 went below 6 times. When 41 per cent of the poor workmen can be eliminated at the expense of only 3 per cent of the good workmen we have a practical tool that points toward better personnel procedures and greater

net profit.

Table 2. Critical score efficiency in prediction.

Test or subtest	:No. "good workmen" :below critical score	: No. "poor workmen" : below critical score
MacQuarrie #1	2	11
MacQuarrie #2	2	19
MacQuarrie #3	0	13
MacQuarrie #4	1	11
MacQuarrie #5	0	17
MacQuarrie #6	2	15
MacQuarrie #7	1	16
Personnel Test	0	13
Total	8	115

"Poor Workmen" Might be Successful in Other Departments. Logically, some of the workmen in the "poor workmen" group could have succeeded in other departments than the one in which they failed. A study of the "poor workmen" was made to see if they could "fit" into other departmental "profiles" in the light of their scores. In some of these cases their scores only barely met the requirements of some other department where their weakness would be less apparent; in others the prediction seemed good. Out of a group of less than fifty who were tried out in other departments, two-thirds were estimated to have done passable work, although the study was inconclusive because of limited time and lack of accurate follow-up data, coupled with

the fact that the study was made during a time when recruiting was difficult. The correlation studies tend to bear out the conclusions of this limited study.

Uniformly High Scores Not Required. It should be mentioned in passing that few of the departments required uniformly high scores or profiles on the MacQuarrie. Tool and die workers and inspectors seemed to require the most generally high profiles. In the special departmental studies it developed that stamp press operators required a quotient greater than unity when the standard score for Test 1 was divided by the standard score for Test 2. The coefficient of validity between this quotient and actual production was .60 and when added to three other measures in a multiple correlation, produced a multiple coefficient of validity of .82. The multiple regression formula included a personality scale which had an independent validity coefficient of .66.

Intercorrelations and Factor Analysis

The foregoing information concerning critical scores, significant differences in means, and the nature of the distributions properly prefaces a discussion of the factors underlying the score relationships.

Characteristic of Sample in Relation to Factor Analysis. Because the literature contains factor analyses of the MacQuarrie test which show various weaknesses as to heterogeneity of samples

(and heterogeneity always produces higher correlations)[2] the author felt that the construction of two factor matrices for these two groups would reveal important differences between them. Workers in these groups are homogeneous in the fact that they are production workmen from one plant and that there is little to quarrel with in the basis for having placed them into these groups. On the other hand, they are quite heterogeneous in age (age range in "good workmen" is from 19 to 58, in "poor workmen" from 20 to 56) and quite heterogeneous in the type of work they do, (ranging from simpler tasks to more complicated tasks). It must be pointed out, however, that the work done by the workmen in this study was far less heterogeneous in character than that of studies in the literature which included students, and clients of a counselling service. The author was seeking general factors rather than specific validation in this study. To discover such general factors requires a sample which is homogeneous in character on the side of the criterion, but heterogeneous through a range within the universe to which the general information might be applied.[3]

MacQuarrie "Total" Score. It is to be noted that in the descriptions of the distributions no mention was made of the MacQuarrie total score. The author's experience with this particular combination of scores of the subtests has been dis-

---

[2] Donald E. Super, "Appraising Vocational Fitness", p. 267.

[3] L. L. Thurstone, "Multiple Factor Analysis", p. 325.

couraging because the score produced little or no validity that could not be better appraised by single subtests or simpler combinations of subtests when taken alone. To economize in effort, then, this score was not used in any of the correlations.

Legend for the Correlation Matrices. Table 3 shows the matrix of the intercorrelations for the "good workmen" and the "poor workmen". In these matrices, Tests 1 through 7 are the seven subtests of the MacQuarrie and Test 8 is the Personnel Test. This legend for the tables will continue throughout the subsequent tables of the factor matrices.

Significant Differences in Correlation Coefficients. In the matrix of the "poor workmen" certain of the correlation coefficients in the upper right-hand half of the matrix are shown underlined. These are the coefficients which differ by more than four times their probable error from the coefficients in the same cell in the matrix of the "good workmen". These differences are assuredly significant and they number 15 out of the 28 coefficients in the matrix. A total of 26 of the coefficients are higher in the "poor" group than in the "good" group. Fourteen of the coefficients so underlined are higher in the "poor" group than in the "good" group. The one which is lower (by the difference between .25 and .47) in the "poor" group is the correlation between Test 6 (Blocks) and Test 2 (Tapping) and it is to be remembered that Test 6 was the distribution which, in the "poor" group, was badly skewed and piled up 9 scores at zero. This would have reduced the coefficient and it cannot be regarded

Table 3. Correlation matrices.

	Tests							
	1	2	3	4	5	6	7	8
				103 Good Workmen				
1	--	.30	.52	.47	.23	.21	.39	.23
2	.30	--	.46	.28	.18	.47	.26	.28
3	.52	.46	--	.33	.29	.25	.35	.15
4	.47	.28	.33	--	.51	.53	.40	.43
5	.23	.18	.29	.51	--	.48	.39	.43
6	.21	.47	.25	.53	.48	--	.52	.61
7	.39	.26	.35	.40	.39	.52	--	.33
8	.23	.28	.15	.43	.43	.61	.33	--
				104 Poor Workmen				
1	--	.49	.40	.65	.55	.56	.49	.38
2	.49	--	.67	.38	.44	.25	.40	.27
3	.40	.67	--	.54	.57	.31	.48	.53
4	.65	.38	.54	--	.73	.63	.63	.49
5	.55	.44	.57	.73	--	.54	.67	.64
6	.56	.25	.31	.63	.54	--	.69	.53
7	.49	.40	.48	.63	.67	.69	--	.66
8	.38	.27	.53	.49	.64	.53	.66	--

Legend for Table:

1. The MacQuarrie Test 1 is a test of steadiness of eye-hand-arm-finger dexterity.
2. The MacQuarrie Test 3 is a test of speed with minimum accuracy, a tapping test.
3. The MacQuarrie Test 3 is a test of aiming, mainly eye-hand-arm coordination.
4. The MacQuarrie Test 4 is a test of two-dimensional spatial relations; copying.
5. The MacQuarrie Test 5 is a test of two-dimensional visualization, changes in size.
6. The MacQuarrie Test 6 is a test of 3-dimensional perception, blocks.
7. The MacQuarrie Test 7 is a complicated curving-line maze of stressing visual acuity.
8. The Personnel Test (Wonderlic) is a spiral-omnibus mental efficiency test.

as a reliable difference, although it contains some illuminating conclusions which will be discussed below.

It was expected by the author that the coefficients of correlation in the matrix for the "poor workmen" would be slightly higher because the standard deviations were slightly greater. But these greater standard deviations in no way account for the size of the differences expressed in the actual obtained coefficients.

Both of the matrices tend to refute conclusions expressed by Babcock and Emerson[4] toward the generalization that the MacQuarrie Test is measuring intelligence since, in the "good workmen" the highest single correlation with The Personnel Test (8) is, .61, the next two highest are .43, and the remainder .33, .28, .23, and .15 in descending order. The author feels that it is safe to say that Tests 1, 2, and 3 are <u>not</u> measuring intelligence as it is commonly defined and that it is doubtful if Tests 4, 5, and 7 have any higher relationship with intelligence (or mental efficiency) than might be required (a) To understand verbal test instructions or, (b) in the more practical situation, to understand training instructions required of any skilled production worker.

It must be pointed out in passing that where self-administering instructions are used with the MacQuarrie there would in-

---

[4] Harriet Babcock and Marian Reves Emerson, "An Analytical Study of the MacQuarrie Test of Mechanical Ability", <u>Jour. of Ed.</u> Psych., Vol. 29, Jan., 1938, pp. 50-55.

evitably be the introduction of another factor--that of reading speed and comprehension. Associated Personnel Technicians, Inc., used the form of the MacQuarrie which deletes the self-administering printed instructions and had monitors to assure that instructions were understood after the practice test had been given. It is quite probable that the high validity coefficients which have been obtained in studies of small, homogeneous groups of workers can be traced to this careful administration method. Where the printed self-administering instructions were before them to distract them from the test administrator's instructions, or where they were allowed to read them by themselves, it is believed that the scores of subjects taking the test would correlate much higher with any spiral omnibus test of mental efficiency or intelligence.

It is, of course, quite probable that the higher correlations of The Personnel Test (8) with the MacQuarrie Tests in the "poor workmen" group are traceable to a reduced capacity to understand the instructions. The correlations are not so high as to dictate that the MacQuarrie scores are governed by intelligence scores, but they do clearly indicate that a critical score for The Personnel Test might well be used prior to additional testing. That score would be quite low, however, since the mean of even the "good workmen" stood at the 36th percentile of a random sample of 7,090 applicants for work in the same community. The mean of the poor workmen was even lower and the bi-serial coefficient of validity between the poor workmen and the good workmen was .32.

Another rather interesting observation can be drawn from these matrices. Note that the correlation of Test 4 with 5 is .73 in the "poor workmen" and .51 in the "good workmen". While the bi-modality of the distribution on Test 4 in the "poor workmen" may contribute to this difference, there is also the possibility that certain persons not destined to succeed at production work may lack the trait of being able to see changes in size in two-dimensional spatial relations and that, as a result, the "poor workmen" were limited in their performance on Test 5 to the simple, two-dimensional spatial relationship needed for Test 4. Also it must be remembered that the bi-modality of the distribution on Test 4 which raised the coefficient, could not of itself have brought it to a point as high as .73. Nor does the bi-modality argue against using the coefficient in a factor matrix.[5]

Hypothesis and Corollary. The author believes that these two matrices justify the following hypothesis and its corollary:

Hypothesis. Since correlations are lower in the matrix for the "good workmen" there is less generalized aptitude and there are greater specific aptitudes in the "good workmen".

Corollary. Since correlations are higher in the matrix for the "poor workmen" there is less specific aptitude and the matrix

---

[5] L. L. Thurstone says, in "Multiple Factor Analysis", page 325: "No assumption of normality of the distribution is involved in factorial analysis. If we could find bimodal distributions for some tests and factors, they would be especially interesting, provided, of course, that the bi-modalities were not artifacts." This bi-modal distribution, then, can only be criticized on its origin and it is certainly not an artifact insofar as this population of "poor workmen" is concerned.

is more heavily loaded with some general factor or factors retarding all of the more specific aptitudes measured.

The hypothesis expressed calls for a multiple-factor analysis to see if it can be determined what such general factors might be. To properly understand the factors extracted requires that the multiple factor analysis be made of both matrices.

In defense of the hypothesis words "greater specific and less generalized aptitudes" in the good workmen, the author points out that his other studies within the same organization tend to show sharp profile differences and changes from department to department and task to task. Specific validity studies may emphasize the need for a high score on one subtest or several, but usually only two or three are emphasized in any ordinary department or task. Thus when a group of workers from different departments is brought together experimentally, the intercorrelations should be lower if they are succeeding, higher if they are failing. The author attributes considerable significance to this fact since it bears out the fundamental concept of differential psychology as it applies to better use of the labor force through improved placement.

The Multiple-Factor Analysis. Since we have been warned with respect to speculation about the nature of factors in a Multiple-factor Analysis,[6] it seemed best to inspect the tables of factor

---

[6] C. C. Peters and Walter R. Van Voorhis, "Statistical Procedures and Their Mathematical Bases", p. 276.

loadings in the light of the foregoing data about the samples and the differences in the distributions.

Table 4 shows the factor loadings of each of the 6 factors which were extracted before the residuals vanished and sets them out side by side for the "good workmen" and the "poor workmen". Roman numerals at the top of the columns are the factors in the order in which they appeared; standard numerals at the left of each row are the same as those in the correlation matrices, indicating from 1 through 7 the seven subtests of the MacQuarrie Series with 8 denoting The Personnel Test; the sub-divided columns under the Roman numerals represent, where headed "G", the loadings for the "good workmen" and where headed "P", the same loadings for the "poor workmen".

Factor I. Since these loadings were heaviest in the "poor workmen" and less so in the "good workmen", we interpret this factor in the light of the previously-expressed hypothesis and corollary: This must be the factor which to a greater degree than the others, is retarding the performance of the "poor workmen" and not retarding the performance of the "good workmen". Where this factor does not retard the performance of the subject, other individual differences come into greater prominence and we have lower intercorrelations (as in the "good workmen"). But where, in the tested subject, this factor retards performance, it serves as a common limiting factor which increases the size of the intercorrelations (as in the "poor workmen"). Since, in the "good workmen", it shows its highest loadings in Test 6, it must

Table 4. Contrast in factor loadings (factor loadings have not been rotated).*

	Factors											
	I		II		III		IV		V		VI	
Tests	G**	P***	G	P	G	P	G	P	G	P	G	P
1	.58	.71	.34	-.10	.26	-.33	-.18	.09	-.06	.14	.13	.04
2	.54	.61	.27	-.56	-.36	.13	.14	.22	.12	.12	.09	.14
3	.58	.71	.48	-.34	.12	.29	.20	-.13	.12	-.06	-.05	.19
4	.70	.81	-.12	.10	.18	-.22	-.15	-.16	.17	.14	.16	.15
5	.60	.83	-.29	.13	.20	.09	.14	-.20	.18	.20	-.15	-.07
6	.74	.71	-.35	.31	-.30	-.22	.17	.20	.05	-.14	.20	.07
7	.63	.80	-.05	.28	.11	.12	.14	.20	-.29	-.07	.10	.11
8	.62	.71	-.34	.24	-.21	.27	-.20	-.09	.05	.22	-.15	-.09

* Since these factor loadings are from contrasting matrices rather than a single matrix, the author felt it unnecessary and undesirable to rotate the axes lest the contrasts become less apparent in the process. This opinion was confirmed by T. L. Kelley, Ph.D., authority on this technique in interpreting validation data.

** G is Good Workmen.

*** P is Poor Workmen.

have much to do with the ability to imagine (or "see") perspective drawings as representations of solids. This cannot be its only characteristic, however, since all of the visualization tests (4, 5, 6, 7) are heavily loaded and so is The Personnel Test. There are good arguments for naming it "Concrete Visual Imagery" since several of the items of The Personnel Test call for this type of imagination for their rapid solution and since even Tests 1, 2, and 3 of the MacQuarrie are so constructed that a right-handed person must estimate what is beneath his hand to his right as he works, rapidly, on every other row of the record test for Tests 1 and 3 and for every row of Test 2.

Glancing at the loadings for "poor workmen" we are struck instantly with the fact that Test 2 has the smallest loading and that it was the one test which was completely uniform in spacing and requires no estimation at all in progressing from point to point.

While there are those who will consider that this interpretation of Factor I constitutes a criticism of the MacQuarrie Test as a practical instrument, the author cannot agree. The same factor of visual imagery is continually required in most modern production work. One does not look toward the pile of raw steel each time he picks up a sheet to be put into the stamp press to be formed, nor does one stack the formed piece deliberately and with his eye on the stack. To do this would slow the operator down to the point at which present-day costs could not be borne competitively. Similarly, in each type of production work repre-

sented in this study, there is that aptitude or trait required, that of "seeing" without "looking", often with that "seeing" task outside the range of peripheral vision as well as focal vision.

The author feels that the faculty of "seeing" without "looking" is the real determinant in this factor and that it is the greatest single individual difference to be found between those for whom success at this type of work can be predicted and those for whom failure is the most certain estimate.

Factor II. Here we have a factor which clearly focuses on the use of one's hands, eyes, and arms together; a psycho-motor efficiency factor. This has been called by other researchers "Manual Dexterity" and that designation as to its nature might well be repeated here were it not for the fact of the differences between the two groups in their factor loadings. It was evident that the "poor workmen" were slower since the greatest difference between their scores and those of the "good workmen" occurred on the Tapping test (Test 2), the purest test of rhythmic speed, but there was almost as great a difference in the next test, Test 3, where accuracy and aiming come into play. The factor included some strong element of control as well as speed, some molar aspect of nervous system functioning as well as pure muscularity. That this is a very important difference between the two groups was evident from the fact that the signs were reversed in each test in each group. That it was more involved with motor activity than with intellective functioning was evident from the fact

that the signs also reverse, in both groups, as between Tests 1, 2, 3, and 4, 5, 6, 7, and 8. The author considers that this factor is "Psycho-Motor Efficiency".

Factor III. This seems to be a measurement of the degree to which the individual places accuracy above speed or the converse. The "poor workmen" seemed to work as rapidly as possible, letting the chips fall where they may, expecting to compensate, in speed, for the errors they make. The "good workmen" seemed to rate their speed in terms of their accuracy. To substantiate this hypothesis, we observe that the test requiring the least accuracy was loaded negatively in the "good workmen" and positively in the "poor workmen"--Test 2, a simple tapping test. Similarly, the test which slowed the hyper-accurate person down the most was the first test (passing the pencil through the narrow openings), and here the factor negatively loaded the "poor workmen" matrix but loaded the "good workmen" matrix positively. This test shows the greatest difference, too, in the factor loadings of Factor III, ranging .59 through .00 from plus .26 to minus .33.

Test 4, too, shows the same factor loading relationship as Test 1.

Arguing against this hypothesis is the relationship of the loadings on The Personnel Test, where the "good workmen" had a negative loading and the "poor workmen" a positive one. The author's explanation for this is that The Personnel Test was a test which did offer some advantage to the person who gambled an

answer in the interests of speed. The trait which we have here described hypothetically as best fitting these data could have the effect of improving somewhat the scores of the "poor workmen" on The Personnel Test and might lower the scores of the "good workmen".

This hypothesis is further supported by the study mentioned earlier of a group of stamp press operators, where the standard score for Test 1 was divided by the standard score for Test 2 to arrive at a quotient which correlated .60 with actual production. Stated simply, this means that within the range of the scores made by those stamp press operators, it is more important that Test 1 exceed Test 2 than that either be high of themselves. A quotient of unity was the average point, as in any normal test quotient, and quotients greater than unity predicted good production.

It is difficult to describe the factor in everyday terms. It could be named "Accuracy Versus Speed", or, with less certainty "Careful Work Habits". More technically it is the degree to which the person adapts his speed to the accuracy which he can produce. It tends to separate the wasteful workers from the ones who conserve materials. If the tendency could be "trained out" of typical workmen who showed the tendency to work faster than their accuracy would permit, then it could logically be called "Careful Work Habits".

Factor IV. This may be the "Perception of Detail" factor

mentioned by Harrell,[7] though this author would prefer to call it "Attention to Detail". The loadings of Tests 1, 3, and 5 prompted this conclusion, since these showed the greatest differences in loadings between the two experimental groups and were the tests requiring the most careful attention to spatial tolerance in execution by the subject. The data are not conclusive. The factor seems clearly to differentiate, but the author prefers to let each observer draw his own conclusions from the data.

Factor V. This seems to be best called "visual acuity" as expressed especially in situations where the eyes must move rapidly from one situation to another. This conclusion comes about from observation of the fact that the loadings (without regard to sign) were heaviest in the "good workmen" in Tests 7, 5, and 4 in that order and that greatest differences between the two experimental groups are expressed in Tests 1, 3, 6, 7, and 8. Again the data are not conclusive.

Factor VI. This seems to show no consistency justifying an explanation of the loadings. An argument could be erected claiming that this factor is "Abstract Intelligence" (in this case, the lack of it because the significant loadings were negative). The Personnel Test and Test 5 both have negative loadings and these two tests introduce the strongest elements of abstract reasoning (although even in Test 5 this was limited to location

---

[7] Willard Harrell, "Validity of Certain Mechanical Ability Tests for Selecting Cotton Mill Machine Fixers", Jour. Soc. Psych., Vol. 8, May, 1937, pp. 279-282.

of a dot in a small square in relation to a letter of the alphabet in a square 4 times as large). Conversely, while Test 6 placed a premium on three-dimensional visualization and was heavily loaded in the "good workmen" it had no element of the abstract in it, the blocks being visible in perspective and requiring no abstract thinking to arrive at the correct answer. Again, the author prefers to avoid any definite conclusion.

USE OF THE FINDINGS

The Industrial Scene

Use of the MacQuarrie Test for Selection. These findings support the research of others who have found the MacQuarrie useful in various industrial situations for improving the selection of production workers. Since most of the studies were for homogeneous groups of workers performing one task, and including a full range of performance in the criterion, this study extends the value of the test to situations in which it is desirable to find out first if the candidate for employment should be included in the total cadre of production workers at all, then to determine the more exact locus of his most productive activity. The whole profile may be inspected for this purpose before studying the details of the relationship of the seven subtests to seek the best possible placement of the candidate.

In addition, the findings with respect to "critical scores"

reveal the value of this instrument for the production concern since it clearly eliminates 41 per cent of the "poor workmen" at a risk of eliminating only 3 per cent of potentially "good workmen".

Specific validation, similar to some of the work done for The Coleman Company, is needed before the personnel technician can use the test for placement. However, the literature is replete with studies of specific tasks to aid the personnel technician in carrying out this secondary phase of the interpretation.

Use of the MacQuarrie for Improving Placement. A glance at a profile (where all of the scores lie above the "critical" minimums) can frequently mislead the placement worker when one or more scores dip quite low (but not below the critical minimums). It becomes quite evident from this study that such ups and downs are to be expected rather than deplored since it is characteristic of "good workmen" to show these wide differences (as demonstrated by lower intercorrelations in the "good workmen" than in the "poor workmen"). This leaves only the problem of considering those persons with generally-high scores (sometimes called the "high flat" profile). Are they "poor workmen" because their scores are together in high zones? This seems to the author to be answered best in this manner: The high profiles are descriptive of those persons who should be considered for higher-level tasks not of the type associated with pure production work. The author's studies of tool and die makers indicated that they would

have high scores on all but the Tapping test. Furthermore, studies by the author and by Irvin T. Shultz tend to support the hypothesis that uniformly high scores tend to be descriptive of engineers.

The generalization may be made that very low single scores (below the "critical" minimum) or scores which are in the middle and low zones which make a uniformly "flat" profile are descriptive of those persons whose futures do not lie in production work; that scores which vary up and down, some high and some low are descriptive of good production workmen if placed properly and according to their best measured talents as shown by the tests, and finally, that those who are uniformly high are destined for tasks above the level of ordinary production work or in skills standing at the peak of a scale of production skill.

The Guidance and Counseling Area

The author feels that these findings materially improve the use of the MacQuarrie for vocational guidance and counseling. Entirely apart from the fact that low scores on Test 7 were quite diagnostic of the need for examination of the subject's vision, there are other useful applications of the test in these fields. The comments under "placement" (above) can be used to make a general orientation of the individual's proper placement in the world of work before using longer and more specialized equipment tests to make more refined judgments.

Where school counselors of late adolescents desire an instrument to direct some students toward industrial arts they may find the "good workmen" pattern, when interpreted along with scores on general mental efficiency, diagnostic of a vector in the direction of production work. The uniformly high scores, coupled with mental efficiency tests that indicate reasonable ceilings for advanced education, tend to indicate vectors toward engineering and other skills requiring good visualization and manual accuracy. The counselor, by subtraction, has a remaining group who score high on the mental efficiency measures, but who have low scores on the MacQuarrie. For them the counselor needs to explore the communicative and uplift pursuits and make appropriate recommendations in those lines based on his findings from other measuring instruments.

It is not the author's intention to claim from this that the MacQuarrie is the answer to the problem of the guidance worker or counselor. Rather, it is a convenient, economical, and efficient tool for preliminary orientation and evaluation.

SUMMARY AND CONCLUSIONS

The MacQuarrie Mechanical Performance Ability Series is an instrument which will aid the personnel worker in discriminating between those workers destined to succeed at production work and those destined to fail.

Those destined to succeed in some department or other based

on the matching of their profile for the seven subtests with departmental validated profiles, will reflect greater variability through the profile than those destined to fail, since the correlations are almost uniformly higher in the matrix of "poor workmen" than in the matrix of "good workmen". "Good workmen" are more "different" throughout a plant; "poor workmen" are more "alike" on this test, the "alikeness" stemming from common factors which cause them to fail.

The factors prompting failure in the "poor workmen" are:

1. Inability to erect concrete visual images of work outside the range of focal vision and even peripheral vision.

2. Lack of psycho-motor efficiency, an attribute of the whole nerve-muscle (and possibly vascular) system.

3. Tendency to work more rapidly than they can work within the limits of accuracy.

Other factors appear to exist, but data are inconclusive for their description or positive identification.

Critical scores on the MacQuarrie Mechanical Performance Series tended to eliminate 41 per cent of the workmen for whom failure occurred at the expense of less than 3 per cent of the workmen rated "good". The remainder of the "poor workmen" have a chance at success in some department where their profiles more clearly match the departmental profile than the departmental profile of the department in which they failed.

The seven subtests of the MacQuarrie and The Personnel Test showed significant differences in mean scores of the two groups

in excess of 999 chances in 1,000 (except for 1 test in which chances were only 997 in 1,000) that the differences were greater than zero.

Conclusions from the data are limited by a difference of 1.3 years in the average education of the two groups and the fact that there were approximately 5 per cent more women in the "good workmen" sample. Age average and range of age was comparable and experience was beyond the calculated learning period in each department, although the "good workmen" had a higher average experience with the company (33.2 months against 24.8 months). In this connection it should be mentioned that save for the fact that the sample was collected toward the close of the war under conditions of a limited labor force, it is doubtful if a sample of "failing workmen" of that size could be collected for experimental purposes.

Considering the amount of data obtained from its administration and the short time required, the author believes the MacQuarrie Mechanical Performance Series to be the most economical means to screen and place factory production workers when coupled with three other brief tests in a battery. It yields more information for the effort and money expended than could be obtained with much more testing. Careful administration is required and accurate collecting of validation data must be a preliminary to any such use of the test.

ACKNOWLEDGMENTS

The author is grateful, indeed, for the instruction, guidance, and inspiration of J. C. Peterson, Ph.D., major professor under whom he studied at Kansas State College and to Cecil B. Read, Ph.D., professor of Mathematics at the University of Wichita who was the author's instructor in Statistics and who lent encouragement and gave counsel as to method.

To Clarence Coleman, Vice President and Assistant General Manager of The Coleman Company, Inc., Floyd D. Quinlisk, Personnel Director of the same company, and M. B. Goff, Employment Manager, the author expresses his appreciation for having encouraged the study and for having provided a good scientific atmosphere for administering the tests and obtaining the ratings.

It was the late Edwina A. Cowan, Ph.D., who first revealed to the author some of the subtleties of the MacQuarrie test and he takes this opportunity to pay tribute to her memory.

Irvin T. Shultz, Ph.D., then of Friends University and now of Baylor University, supervised the author's early excursions into measurement of behavior and encouraged thoroughness of method in the first phases of the study.

Finally, but not in any measure least deserving of mention, is the author's wife, Pearl R. Barnabas, to whom he is indebted not only for inspiration, but for her aid as an assistant in thorough test administration in the first phase of the study and for the contributions of an exceptionally-orderly mind throughout the progress of the analysis.

REFERENCES

Introductory Comments

The author has included in these References those studies in the field which tend to relate to the problem under discussion in this Thesis. Thus the references divide logically into the literature relating to: Production tasks in which the validity of the MacQuarrie was explored; the factor analyses that have been made of the test; the general literature of factor analysis; and general reference material on both the MacQuarrie and The Personnel Test which was included in the battery for this study and in the intercorrelations.

The studies of Rose G. Anderson, Milton Blum, Robert L. Chapman, Edwin Ghiselli, Charles Goodman, Willard Harrell, and Irvin T. Shultz relate to production tasks. In addition, the studies of Goodman, Chapman, and Harrell have to do with factor analyses made of a single matrix of the test. Other comments with respect to use of the test in factor analysis are found in L. L. Thurstone's "Primary Mental Abilities". The general literature of factor analysis is discussed in the works of Thurstone that are listed, along with the book written by Truman L. Kelley, the Peters and Van Voorhis' "Statistical Procedures and Their Mathematical Bases", and "The Factorial Analysis of Human Ability" by Thomson. The remainder of the material listed is general reference material which will be useful to the person who sets out to

explore the worth of the test in other situations.

In connection with the study of Babcock and Emerson, it is interesting to point out that in the light of this study their conclusions tend to show that the sample from which they derived their data was a sample of "poor workmen" (in the language of this Thesis). This is quite logical, since their sample was of cases coming to the New York Vocational Adjustment Bureau, cases which would include numbers of persons who needed a new occupational vector. They found higher intercorrelations in the subtests and higher correlations with their intelligence scores. Thus while they were careful to eliminate those persons suffering from other disorders, they had no objective basis for dividing their sample into palpably good workmen or palpably poor workmen based on actual performance on the job.

The author found no studies in the literature in which contrasting matrices were constructed for validation purposes. This seems to be a useful tool for discovering the underlying reasons for valid contrasts in individual differences in subjects grouped contrastingly and is recommended for the investigation of those who are doing work in the field. While the construction of multiple coefficients of correlation will produce weights when brought to bear upon a dependent variable, and will show the comparative value of a series of tests in prediction, separate matrices bring out into sharp relief the general factors underlying (and perhaps causing) the differences measured.

Periodicals

Anderson, Rose G. "Test Scores and Efficiency Rating of Machinists," Journal of Applied Psychology, Vol. 31, 1947, pp. 377-388.

Babcock, Harriet, and Marion Reves. "An Analytical Study of the MacQuarrie Test for Mechanical Ability," Journal of Educational Psychology, Vol. 29, January, 1938, pp. 50-55.

Bingham, Walter V. "MacQuarrie Test for Mechanical Ability," Occupations, Vol. 14, December, 1935, pp. 202-205.

Blum, Milton. "Selection of Sewing Machine Operators," Journal of Applied Psychology, Vol. 27, February, 1943, pp. 35-40.

Chapman, Robert L. "The MacQuarrie Test for Mechanical Ability," Psychometrika, Vol. 13, No. 3, September, 1948, pp. 175-179.

Chiselli, Edwin. "Tests for the Selection of Inspector-Packers," Psychological Bulletin, Vol. 38, October, 1941, p. 735.

Goodman, Charles H. "MacQuarrie Test for Mechanical Ability: II. Factor Analysis," Journal of Applied Psychology, Vol. 31, No. 2, April, 1947, pp. 502-510.

Harrell, Willard. "A Factor Analysis of Mechanical Ability Tests," Psychometrika, Vol. 5, No. 1, March, 1940, pp. 17-33.

MacQuarrie, T. W. "A Mechanical Ability Test," Journal of Personnel Research, Vol. 5, No. 9, January, 1927, pp. 329-337.

Shultz, Irvin T. "Comparison of the Scores of Design Engineering Personnel with Factory Workers on the MacQuarrie Mechanical Performance Ability Tests," Transactions Kansas Academy of Science, Vol. 49, No. 3, 1946.

Books

Kelley, Truman L. Crossroads in the Mind of Man. Leland Stanford University: Stanford University Press, 1928.

Peters, C. C., and Walter R. Van Voorhis. Statistical Procedures and Their Mathematical Bases. New York: McGraw-Hill, 1940,

Stead, William H., and Carroll L. Shartle. Occupational Counseling Techniques. New York: The American Book Company, 1940, pp. 96-98, 102, 103.

Super, Donald E. Appraising Vocational Fitness by Means of Psychological Tests. New York: Harper and Brothers, 1949.

Thomson, Sir Godfrey H. The Factorial Analysis of Human Ability. New York: Houghton Mifflin Company, 1939.

Thurstone, L. L. The Vectors of Mind. Chicago: University of Chicago Press, 1935.

_____ Primary Mental Abilities. Chicago: University of Chicago Press, 1938, pp. 31, 41, 42.

_____ Multiple-Factor Analysis. Chicago: University of Chicago Press, 1949.

Tiffin, Joseph. Industrial Psychology. New York: Prentice Hall, 1947, pp. 129-130.

A STUDY OF CERTAIN GENERAL FACTORS PREDICTING SUCCESS AT PRODUCTION WORK

by

BENTLEY AUGUSTINE BARNABAS

A. B., University of Wichita, 1929

AN ABSTRACT OF A THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Psychology

KANSAS STATE COLLEGE
OF AGRICULTURE AND APPLIED SCIENCE

1953

The author obtained ratings of 103 "good workmen" and 104 "poor workmen" throughout the production departments of The Coleman Company, Incorporated, of Wichita, Kansas, for the purpose of developing critical scores on the MacQuarrie Test of Mechanical Ability and the Personnel Test to differentiate the two types on a plant-wide basis and to discover any general factors which might underly success at production work.

Distributions were made of the scores of both groups, critical scores calculated and, where indicated, bi-serial r's; two intercorrelation tables were constructed and a factor analysis was made of the matrices.

The author found significant differences in the means of all seven of the MacQuarrie Tests and the Personnel Test, ranging from 2.81 (in the quotient of the actual difference divided by the standard error of the difference) for the MacQuarrie Test 1 to 4.48 for Test 7, with 3.76 for the Personnel Test. All were above 999 chances in 1,000 that the difference is greater than zero except for Test 1 which was 997 chances in 1,000. Critical scores revealed that 41 per cent of the poor workmen could have been eliminated with a loss of only 3 per cent of the good workmen had the critical scores been used to predict.

Intercorrelation tables revealed significant differences four times the probable error of the r's in 15 of the 28 correlations, the r's running higher in the "poor workmen". Factor loadings, unrotated, showed Factor I which was named "Concrete Visual Imagery", Factor II "Psycho-Motor Efficiency", and

Factor III "Accuracy Versus Speed", all showing differences in the two groups. Factors IV, V, and VI were difficult to name and the author limited opinions as to their possible meaning.

The author concluded that the MacQuarrie Test is very useful in that it shows greater variation of individual differences as measured by different subtest scores in the "good workmen" than the "poor workmen" and reveals little correlation with the Personnel Test. He further concluded that the test can be very useful in eliminating those not destined for success at production work through the use of critical scores, can be used to make a better placement of workers who might fail in one department but who could succeed at another task and that workers in the areas of guidance and counseling can use it to direct workers toward manual skills, mental skills, or mental-manual skills by comparing the MacQuarrie scores with mental efficiency scores and using this as a guide to further exploration.